



## SEQUENCE LISTING

<110> Kosan Biosciences, Inc.  
Julien, Bryan

<120> TRANSFORMATION SYSTEM BASED ON THE  
INTEGRASE GENE AND ATTACHMENT SITE FOR MYXOCOCCUS XANTHUS  
BACTERIOPHAGE MX9

<130> 300622009940

<140> US 10/645,818

<141> 2003-08-20

<150> US 60/405,196

<151> 2002-08-21

<160> 20

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1647

<212> DNA

<213> Bacteriophage MX9

<400> 1

```
gtggcgctca ggggtgcgtc ggacgccact accaaccctt ctcgacttgt gcagtccgtc 60
gccgccggcc cgcgtgcgac tccgtgggtt gtcagtgcgt cgtggtacct gctagggcgt 120
acagcaacgg gggagtacat cgtgagtagc gacgcggcga agaagggcca tccaatggca 180
actgcggcgg agcggttgcc gacgtcacca atcgacgtca acgctctggc gctggagggtg 240
gcccggttg tggccctcca gcagcaaagt gcgacgccgc catcgtccgg ccgcactttc 300
ggcgcggttg cggatgactg gtcctacact gaggccaagc gcctcgtgtg ccccgacaat 360
gagcgccgcc atcttcgcca tatggaggcg ctctggggca tgacggatgt ggagctcacg 420
ccgcgcgtcg tgaaggcgca cctggcgga cttctcaagc cagagggggc gctgagcgca 480
gccaccgtca ataaggtgcg ctctaccggc aagcgcatca tcaaggcggc gcaaatcaac 540
ggcgagtggg gcccggtgaa tcctttcggc gtgctcgacc gcgaaaaaga ggcgaaggcc 600
gagcgccctc cgctgacggc agcggagtgc cgggcgggtg tcccgcactt ccgcgcggac 660
cggcgccgcg agtttctctt ccaggctctt ctggggccac gcccgcgga agagaaggcg 720
ctcctcaagg aagatgtgga cgtcgaggcg cgcaccgtca ttttcggcg cagcaatgga 780
cgagacacga caaagacggg acgcgagcgt cgcgtgccgg tgccggatga gttgtggccc 840
gtgctcctcg atgcgatgca ggccagtcgg tctgacctcg ttttccgaa cgcgaagggt 900
gagaggcagc gcgcagacac gaagatgacg cgcgtgctgc gcaactgcgt atccgcggct 960
ggtgtcgtgg tgggctggga ttacatctgc cgcacgcagg gctgcggcta ccgagatgtg 1020
cagtctggtg gcgcgcgcca ggagcgtcgg tgccccgcct gcgacaagcg catgtggggc 1080
agtggtcgcc ccaaaccgcg cgtctggtac gggctccgtc acaccgcggc gacactgcac 1140
aggaaggcgg gctgcgaccc gctcgtcatc aagctcgtgc tggggcatgc ggctgtcgac 1200
accacggacg acgtgtacac gcacctcgac gaggactact gccgcgccga acttaacaag 1260
ttgtcgtga aggccccgcc gccaccacct actcaccagg gaggaagtga cggcgccct 1320
gactcaggac gcaacaccta cggtgaaaga ggcaccatgc acggattggg agatttgca 1380
catcaccggg cgagagcttg ggaagctcgt gctctaccaa ctgagctacc accgcggaac 1440
ttggccgggg gtataccggc gccgctgctg agcgtcaagg acgttgccgc ttcactctca 1500
gtgagcacgg cgaaggtgta ccagctcctc gccgcggcg tectgectac cgtgtgggtg 1560
ggccagtcgc gcccgctcaa gcgtgaggac ctggacgcct acatcgccc cgcgacggcc 1620
accggcggga agcgggggtg caaatga
```

<210> 2

<211> 548

<212> PRT

<213> Bacteriophage MX9

<400> 2

Val	Ala	Leu	Arg	Gly	Ala	Ser	Asp	Ala	Thr	Thr	Asn	Pro	Ser	Arg	Leu
1				5					10					15	
Val	Gln	Ser	Val	Ala	Ala	Gly	Pro	Arg	Ala	Thr	Pro	Trp	Gly	Val	Ser
			20					25					30		
Ala	Ser	Trp	Tyr	Leu	Leu	Gly	Arg	Thr	Ala	Thr	Gly	Glu	Tyr	Ile	Val
		35					40					45			
Ser	Ser	Asp	Ala	Ala	Lys	Lys	Gly	His	Pro	Met	Ala	Thr	Ala	Ala	Glu
	50					55					60				
Arg	Leu	Pro	Thr	Ser	Pro	Ile	Asp	Val	Asn	Ala	Leu	Ala	Leu	Glu	Val
65					70					75					80
Ala	Arg	Leu	Val	Ala	Leu	Gln	Gln	Gln	Ser	Ala	Thr	Pro	Pro	Ser	Ser
				85					90					95	
Gly	Arg	Thr	Phe	Gly	Ala	Val	Ala	Asp	Asp	Trp	Leu	Ile	Thr	Glu	Ala
			100					105					110		
Lys	Arg	Leu	Val	Cys	Pro	Asp	Asn	Glu	Arg	Arg	His	Leu	Arg	His	Met
		115					120					125			
Glu	Ala	Leu	Trp	Gly	Met	Thr	Asp	Val	Glu	Leu	Thr	Pro	Arg	Val	Val
	130					135					140				
Lys	Ala	His	Leu	Ala	Gly	Leu	Leu	Lys	Pro	Glu	Gly	Pro	Leu	Ser	Ala
145					150					155					160
Ala	Thr	Val	Asn	Lys	Val	Arg	Ser	Thr	Gly	Lys	Arg	Ile	Ile	Lys	Ala
				165					170					175	
Ala	Gln	Ile	Asn	Gly	Glu	Trp	Gly	Pro	Val	Asn	Pro	Phe	Gly	Val	Leu
			180					185					190		
Asp	Arg	Glu	Lys	Glu	Ala	Lys	Ala	Glu	Arg	Leu	Thr	Leu	Thr	Ala	Ala
		195					200					205			
Glu	Cys	Arg	Ala	Val	Leu	Pro	His	Phe	Arg	Ala	Asp	Arg	Arg	Arg	Glu
	210					215					220				
Phe	Leu	Phe	Gln	Val	Phe	Leu	Gly	Pro	Arg	Pro	Gly	Glu	Glu	Lys	Ala
225					230					235					240
Leu	Leu	Lys	Glu	Asp	Val	Asp	Val	Glu	Ala	Arg	Thr	Val	Ile	Phe	Arg
				245					250					255	
Arg	Ser	Asn	Gly	Arg	Asp	Thr	Thr	Lys	Thr	Gly	Arg	Glu	Arg	Arg	Val
		260						265					270		
Pro	Val	Pro	Asp	Glu	Leu	Trp	Pro	Val	Leu	Leu	Asp	Ala	Met	Gln	Ala
		275					280					285			
Ser	Pro	Ser	Asp	Leu	Val	Phe	Pro	Asn	Ala	Lys	Gly	Glu	Arg	Gln	Arg
	290					295					300				
Ala	Asp	Thr	Lys	Met	Thr	Arg	Val	Leu	Arg	Thr	Ala	Leu	Ser	Ala	Ala
305					310					315					320
Gly	Val	Val	Val	Gly	Trp	Asp	Tyr	Ile	Cys	Arg	Thr	Gln	Gly	Cys	Gly
				325					330					335	
Tyr	Arg	Asp	Val	Gln	Ser	Gly	Gly	Ala	Arg	Gln	Glu	Arg	Arg	Cys	Pro
			340					345					350		
Ala	Cys	Asp	Lys	Arg	Met	Trp	Ala	Ser	Gly	Arg	Pro	Lys	Pro	Ala	Val
		355					360					365			
Trp	Tyr	Gly	Leu	Arg	His	Thr	Ala	Ala	Thr	Leu	His	Arg	Lys	Ala	Gly
	370					375					380				
Cys	Asp	Pro	Leu	Val	Ile	Lys	Leu	Val	Leu	Gly	His	Ala	Ala	Val	Asp
385					390					395					400
Thr	Thr	Asp	Asp	Val	Tyr	Thr	His	Leu	Asp	Glu	Asp	Tyr	Cys	Arg	Ala
				405					410					415	
Glu	Leu	Asn	Lys	Leu	Ser	Leu	Lys	Ala	Pro	Pro	Pro	Pro	Pro	Thr	His
			420					425					430		
Gln	Gly	Gly	Ser	Asp	Gly	Gly	Pro	Asp	Ser	Gly	Arg	Asn	Thr	Tyr	Gly
		435					440					445			
Glu	Gly	Gly	Thr	Met	His	Gly	Leu	Gly	Asp	Leu	Gln	His	His	Arg	Ala
	450					455					460				

Arg Ala Trp Glu Ala Arg Ala Leu Pro Thr Glu Leu Pro Pro Arg Asn  
 465 470 475 480  
 Leu Ala Gly Gly Ile Pro Ala Pro Leu Leu Ser Val Lys Asp Val Ala  
 485 490 495  
 Ala Ser Leu Ser Val Ser Thr Ala Lys Val Tyr Gln Leu Leu Ala Ala  
 500 505 510  
 Gly Val Leu Pro Thr Val Trp Val Gly Gln Ser Arg Arg Val Lys Arg  
 515 520 525  
 Glu Asp Leu Asp Ala Tyr Ile Ala Arg Ala Thr Ala Thr Gly Gly Lys  
 530 535 540  
 Arg Gly Gly Lys  
 545

<210> 3  
 <211> 360  
 <212> DNA  
 <213> Bacteriophage MX9

<400> 3  
 gtgagctgac ctcaacggtt tggtgggtgg ggagcgggac agcggaccac atggtgccag 60  
 ggcttacggc ttgcacacg gggctgggag atgctgaacg gagcgtccca tgtccacgcg 120  
 atgccgcctg gcttgacat agggattcga aacctcgacc ccgagcttgg gaagctcgtg 180  
 ctctaccaac tgagctacca ccgcaggcga agcaggcgcc aaagtacggg ccgccctgtg 240  
 gcttgctcaac gggaagtgg gtgctactcc gtctcctcga cggtgagctg gtacgagtc 300  
 tggaagtgg actcgcggtt gcgcgcgtcc cggacctcga agaggtagac gcctggctcg 360

<210> 4  
 <211> 360  
 <212> DNA  
 <213> Bacteriophage MX9

<400> 4  
 cgagccgggg acgggagcgg cgggaccggc ttgcgcgcgt ttacagcatc cttgctgcaa 60  
 gacgccccga ggcccgaaaa gacgaaggcc ggagctcccg agtttctcctca aggactaccg 120  
 gccttcatgg gtgagcggcg gaagggattc gaacctcga ccccgagctt gggaagctcg 180  
 tgctctacca actgagctac caccgcaggc gaagcagggc gcaaagtacg ggccgccctg 240  
 tggcttgctca acgggaagtg aggtgctact ccgtctcctc gacgggtgagc tggtagcagt 300  
 cctggaagt ggactcgcgg ttgcgcgcgt cccggacctc gaagaggtag acgcctggct 360

<210> 5  
 <211> 42  
 <212> DNA  
 <213> Bacteriophage MX9

<400> 5  
 gagcttgagg agctcgtgct ctaccaactg agctaccacc gc 42

<210> 6  
 <211> 240  
 <212> DNA  
 <213> Bacteriophage MX9

<400> 6  
 tgccagggtt tacggcttcg cacacggggc tgggcgatgc tgaacggagc gtcccatgtc 60  
 cacgcgatgc cgcttggtt gcacataggg attcgaaacc tcgaccccgga gcttggaag 120  
 ctccgcctcg acccgtccag gcgttatcag ccgttcgcaa acccttactt cgccttgggg 180  
 attccgggccc gggggcctgt ccatccgtcg cagcgggtag caggaggagt cagggggggtt 240

<210> 7  
 <211> 257  
 <212> DNA  
 <213> Bacteriophage MX9

<400> 7  
 cgccaccacc tactcaccag ggaggaagtg acggcggccc tgactcagga cgcaacacct 60  
 acggtgaagg aggcaccatg cacggattgg gagatttgca gcatcaccgg gcgagagctt 120  
 gggaagctcg tgctctacca actgagctac caccgcggaa cttggccggg ggtataccgg 180  
 cgccgctgct gagcgtcaag gacgttgcg cttcactctc agtgagcacg gcgaagggtgt 240  
 accagctcct cgccgcc 257

<210> 8  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 8  
 gaaggaggca ccatgcacgg 20

<210> 9  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 9  
 ctcaactgaga gtgaagccgc 20

<210> 10  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 10  
 cgaggtccgg gacgcgcgca 20

<210> 11  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 11  
 tgccagggct tacggcttc 19

<210> 12  
 <211> 74  
 <212> DNA  
 <213> Myxococcus xanthus

<400> 12  
 gcggugguag cucaguuggu agagcacgag cuuccaagc ucggggucga gguucgaau 60  
 ccuuccgcc gcuc 74

<210> 13  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 13  
 tatcccagca accgccggag 20

<210> 14  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 14  
 cagcacgggt gcagcaac 18

<210> 15  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 15  
 cccaattggc tcagggcagc ggctcatt 28

<210> 16  
 <211> 31  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 16  
 ccccatggcg ctcaggggtg cgtcggacgc c 31

<210> 17  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 17  
 acaggttggc tgataagtcc ccggtct 27

<210> 18

<211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic Construct

<400> 18  
 ggatccagac cggggactta tcagccaacc tgt 33

<210> 19  
 <211> 1047  
 <212> DNA  
 <213> Chrysoperla carnea (Insect)

<400> 19  
 atggaaaaaa aggaatttcg tgttttgata aaatactgtt ttctgaaggg aaaaaataca 60  
 gtggaagcaa aaacttggct tgataatgag tttccggact ctgcccagcagg gaaatcaaca 120  
 ataattgatt ggtatgcaaa attcaagcgt ggtgaaatga gcacggagga cgggtgaacgc 180  
 agtggacgcc cgaaagaggt ggttaccgac gaaaacatca aaaaaatcca caaatgatt 240  
 ttgaatgacc gtaaaatgaa gttgatcgag atagcagagg ccttaaagat atcaaaggaa 300  
 cgtgttggtc atatcattca tcaatatttg gatatgcgga agctctgtgc aaaatgggtg 360  
 ccgcgcgagc tcacatttga ccaaaaacaa caacgtgttg atgattctga gcggtgtttg 420  
 cagctgttaa ctcgtaatac acccgagttt ttccgtcgat atgtgacaat ggatgaaaca 480  
 tggctccatc actacactcc tgagtccaat cgacagtcgg ctgagtgagc agcgaccggt 540  
 gaaccgtctc cgaagcgtgg aaagactcaa aagtcgcgtg gcaaagtaat ggctctgtgt 600  
 tttttcgtat cgcattggaat aattttttatc gattatcttg agaagggaaa aaccatcaac 660  
 agtgactatt atatggcggt attggagcgt ttgaaggctg aaatcgcggc aaaacggccc 720  
 catatgaaga agaaaaaagt gttgttccac caagacaacg caccgtgccca caagtcattg 780  
 agaacgatgg caaaaattca tgaattgggc ttccaattgc ttccccaccc accgtattct 840  
 ccagatctgg ccccccagcga ctttttcttg ttctcagacc tcaaaaggat gctcgcaggg 900  
 aaaaaatttg gctgcaatga agaggtgatc gccgaaactg aggcctatct tgaggcaaaa 960  
 ccgaaggagt actacaaaaa tgggtatcaaa aaattggaag gtcgttataa tcgttgtatc 1020  
 gctcttgaag ggaactatgt tgaataa 1047

<210> 20  
 <211> 348  
 <212> PRT  
 <213> Chrysoperla carnea (Insect)

<400> 20  
 Met Glu Lys Lys Glu Asn Arg Val Leu Ile Lys Tyr Cys Asn Leu Lys  
 1 5 10 15  
 Gly Lys Asn Thr Val Glu Ala Lys Thr Trp Leu Asp Asn Glu Asn Pro  
 20 25 30  
 Asp Ser Ala Pro Gly Lys Ser Thr Ile Ile Asp Trp Tyr Ala Lys Phe  
 35 40 45  
 Lys Arg Gly Glu Met Ser Thr Glu Asp Gly Glu Arg Ser Gly Arg Pro  
 50 55 60  
 Lys Glu Val Val Thr Asp Glu Asn Ile Lys Lys Ile His Lys Met Ile  
 65 70 75 80  
 Leu Asn Asp Arg Lys Met Lys Leu Ile Glu Ile Ala Glu Ala Leu Lys  
 85 90 95  
 Ile Ser Lys Glu Arg Val Gly His Ile Ile His Gln Tyr Leu Asp Met  
 100 105 110  
 Arg Lys Leu Cys Ala Lys Trp Val Pro Arg Glu Leu Thr Asn Asp Gln  
 115 120 125  
 Lys Gln Gln Arg Val Asp Asp Ser Glu Arg Cys Leu Gln Leu Leu Thr  
 130 135 140  
 Arg Asn Thr Pro Glu Asn Phe Arg Arg Tyr Val Thr Met Asp Glu Thr  
 145 150 155 160

Trp	Leu	His	His	Tyr	Thr	Pro	Glu	Ser	Asn	Arg	Gln	Ser	Ala	Glu	Trp
				165					170					175	
Thr	Ala	Thr	Gly	Glu	Pro	Ser	Pro	Lys	Arg	Gly	Lys	Thr	Gln	Lys	Ser
			180					185					190		
Ala	Gly	Lys	Val	Met	Ala	Ser	Val	Asn	Phe	Asp	Ala	His	Gly	Ile	Ile
		195					200					205			
Asn	Ile	Asp	Tyr	Leu	Glu	Lys	Gly	Lys	Thr	Ile	Asn	Ser	Asp	Tyr	Tyr
	210					215					220				
Met	Ala	Leu	Leu	Glu	Arg	Leu	Lys	Val	Glu	Ile	Ala	Ala	Lys	Arg	Pro
225					230					235					240
His	Met	Lys	Lys	Lys	Lys	Val	Leu	Phe	His	Gln	Asp	Asn	Ala	Pro	Cys
				245					250					255	
His	Lys	Ser	Leu	Arg	Thr	Met	Ala	Lys	Ile	His	Glu	Leu	Gly	Phe	Glu
			260					265					270		
Leu	Leu	Pro	His	Pro	Pro	Tyr	Ser	Pro	Asp	Leu	Ala	Pro	Ser	Asp	Asn
		275					280					285			
Phe	Leu	Phe	Ser	Asp	Leu	Lys	Arg	Met	Leu	Ala	Gly	Lys	Lys	Asn	Gly
	290					295					300				
Cys	Asn	Glu	Glu	Val	Ile	Ala	Glu	Thr	Glu	Ala	Tyr	Asn	Glu	Ala	Lys
305					310					315					320
Pro	Lys	Glu	Tyr	Tyr	Gln	Asn	Gly	Ile	Lys	Lys	Leu	Glu	Gly	Arg	Tyr
				325					330					335	
Asn	Arg	Cys	Ile	Ala	Leu	Glu	Gly	Asn	Tyr	Val	Glu				
			340					345							